

# The best mixing time for random walks on trees

Andrew Beveridge\* and Jeanmarie Youngblood†

October 21, 2014

## Abstract

We characterize the extremal structures for mixing walks on trees that start from the most advantageous vertex. Let  $G = (V, E)$  be a tree with stationary distribution  $\pi$ . For a vertex  $v \in V$ , let  $H(v, \pi)$  denote the expected length of an optimal stopping rule from  $v$  to  $\pi$ . The *best mixing time* for  $G$  is  $\min_{v \in V} H(v, \pi)$ . We show that among all trees with  $|V| = n$ , the best mixing time is minimized uniquely by the star. For even  $n$ , the best mixing time is maximized by the uniquely path. Surprising, for odd  $n$ , the best mixing time is maximized uniquely by a path of length  $n - 1$  with a single leaf adjacent to one central vertex.

## 1 Introduction

We resolve the following extremal question for random walks on trees: what tree structure minimizes/maximizes the expected length of an optimal stopping rule to the stationary distribution  $\pi$ , given that we start at the most advantageous vertex? Naturally, the star  $S_n$  is the minimizing tree structure, but the maximization problem has an unexpected twist. The path  $P_n$  is the maximizing structure when  $n$  is even, but for odd  $n$  the maximizing structure is the near-path  $Y_n$  consisting of a path on  $n - 1$  vertices with a single leaf adjacent to one of the two central vertices. We refer to this graph  $Y_n$  as the *wishbone*. This choice of name is suggested by the layout of  $Y_n$  in Figure 4.1(c) below.

A *random walk* on an undirected graph  $G = (V, E)$  consists of a sequence of vertices  $(w_0, w_1, \dots, w_n, \dots)$  such that  $\Pr[w_{t+1} = w \mid w_t = v]$  is  $1/\deg(v)$  if  $(v, w) \in E$  and 0 otherwise. The *hitting time*  $H(v, w)$  from vertex  $v$  to vertex  $w$  is the expected number of steps before a random walk started at  $v$  visits  $w$  for the first time. We also define  $H(v, v) = 0$ , while  $\text{Ret}(v)$  denotes the expected number of steps before a random walk started at  $v$  first returns to  $v$ .

When  $G$  is not bipartite, the distribution of  $w_t$  converges to the *stationary distribution*  $\pi$ , where  $\pi_v = \deg(v)/2|E|$ . Inconveniently, we do not have convergence for bipartite graphs (including trees), but we can rectify this at the cost of doubling the expected length of any walk. Indeed the distribution  $w_t$  converges to  $\pi$  when we follow a *lazy random walk* in which we remain at the current state with probability  $1/2$  at each time step. A *mixing measure*

---

\*Department of Mathematics, Statistics and Computer Science, Macalester College, Saint Paul, MN 55105

†Department of Mathematics, University of Minnesota, Minneapolis, MN 55455

of a graph  $G$  captures the rate of convergence to the stationary distribution  $\pi$ . Let  $\sigma = \sigma_0$  denote our initial distribution and  $\sigma_t$  denote the distribution for the  $t$ -th step of a random walk. For a fixed constant  $0 < \epsilon < 1$ , the (approximate) mixing time of  $G$  is given by  $T_{\text{mix}}(\epsilon) = \max_{\sigma} \min\{T \geq 0 : \|\sigma_t - \pi\| < \epsilon \text{ for all } t \geq T\}$ , where the maximum is taken over all possible starting distributions and  $\|\cdot\|$  is the given metric. This mixing time depends upon the choice of the parameter  $\epsilon$ .

Lovász and Winkler [9] studied a class of parameterless mixing measures by using more sophisticated *stopping rules* to drive the random walk to a desired distribution. Suppose that we are given a *starting distribution*  $\sigma$  and a *target distribution*  $\tau$ . A  $(\sigma, \tau)$  *stopping rule* halts a random walk whose initial state is drawn from  $\sigma$  so that the final state is governed by the distribution  $\tau$ . The *access time*  $H(\sigma, \tau)$  denotes the minimum expected length for a such a  $(\sigma, \tau)$  stopping rule to halt. We say that a stopping rule is *optimal* if it achieves this minimum expected length. Using access times, we have three natural mixing measures, the *mixing time*, the *reset time*, and the *best mixing time* given respectively by

$$T_{\text{mix}} = \max_{v \in V} H(v, \pi), \quad \text{and} \quad T_{\text{reset}} = \sum_{v \in V} \pi_v H(v, \pi), \quad \text{and} \quad T_{\text{bestmix}} = \min_{v \in V} H(v, \pi).$$

These are the worst-case, average-case and best-case mixing measures. These quantities are called *exact mixing measures*, as opposed to approximate mixing measures like  $T_{\text{mix}}(\epsilon)$  above. See [2] for a taxonomy that compares exact and approximate mixing measures for Markov chains. We also note that exact stopping rules converge to  $\pi$  on a bipartite graph, even for non-lazy walks. Indeed, stopping rules can employ randomness in deciding when to stop, which has the same periodicity-breaking effect as using a lazy walk.

It is natural to wonder which graph structures lead to slow or rapid mixing, and the study of extremal graphs for random walks is well-established, cf. [1, 5, 7]. A natural line of inquiry is to restrict our attention to trees. The extremal tree structures on  $n$  vertices for  $T_{\text{mix}}$  and  $T_{\text{forget}}$  were characterized in [4]. Not surprisingly, the unique minimal structure is the star  $S_n = K_{1, n-1}$  and the unique maximal structure is the path.

**Theorem 1.1** ([4]) *If  $G$  is a tree on  $n \geq 3$  vertices then*

$$\frac{3}{2} \leq T_{\text{mix}} \leq \frac{2n^2 - 4n + 3}{6}$$

and

$$1 \leq T_{\text{reset}} \leq \begin{cases} \frac{1}{4}(n^2 - 2n + 2) & \text{if } n \text{ is even,} \\ \frac{1}{4}(n-1)^2 & \text{if } n \text{ is odd.} \end{cases}$$

*In each instance, the lower bound is achieved uniquely by the star  $S_n$  and the upper bound is achieved uniquely by the path  $P_n$ .*

We add to these extremal tree characterizations by studying the best mixing time  $T_{\text{bestmix}}$ .

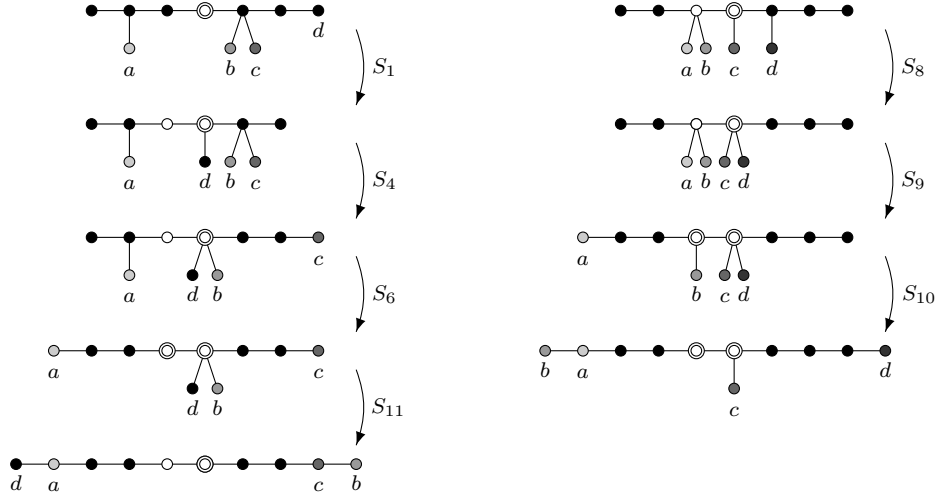


Figure 1.1: Two different trees that are converted to  $P_{10}$  and  $Y_{11}$ , respectively, using the algorithm and the tree surgeries  $S_i$  defined in Section 4. The value of  $T_{\text{bestmix}}$  monotonically increases throughout. The white vertices are the foci, with  $T_{\text{bestmix}}$  achieved by the circled vertex. Certain vertices are labeled by  $a, b, c, d$  so they can be tracked across steps.

**Theorem 1.2** *Let  $G$  be a tree on  $n \geq 3$  vertices. Then*

$$1/2 \leq T_{\text{bestmix}}(G) \leq \begin{cases} \frac{1}{12}(n^2 + 4n - 6) & \text{if } n \text{ is even,} \\ \frac{1}{12}(n^2 + 4n - 15) & \text{if } n \text{ is odd.} \end{cases}$$

*The star  $S_n$  is the unique minimizing tree. For  $n$  even, the path  $P_n$  is the unique maximizing tree. For  $n$  odd, the wishbone  $Y_n$  is the unique maximizing tree.*

Below, we assume that  $n \geq 4$  since there is a unique tree for  $n = 3$  (and the formula is correct in that case). The appearance of the wishbone is quite unexpected. At the end of the next section, we discuss the characteristics of  $Y_n$  that lead to its maximization of the best mixing time for odd  $n$ .

The proof of Theorem 1.2 is more involved than the proof of Theorem 1.1 in [4]. The reason is that a minor change to the tree can abruptly shift the location of the initial vertex achieving  $T_{\text{bestmix}}$ , whereas the behavior of  $T_{\text{mix}}$  and  $T_{\text{reset}}$  are more stable under minor structural alterations. Here is an overview of our proof. First, we replace the given tree with a caterpillar having a larger best mixing time. We then follow a prescribed algorithm, carefully moving one or two leaves at a time via a *tree surgery*  $S_i$ . These surgeries monotonically increase  $T_{\text{bestmix}}$ , until we arrive at an even path or an odd wishbone. Figure 1.1 shows two sequences of surgeries, the first resulting in  $P_{10}$  and the second resulting in  $Y_{11}$ . The surgeries and the guiding algorithm are described in Section 4.

## 2 Preliminaries

For an introduction to the theory of exact stopping rules, see [9]. To prove Theorem 1.2, we need the formula for access times from singleton distributions to the stationary distribution. Therefore, we limit ourselves to describing results from [9] needed to calculate these values.

### 2.1 Pessimal vertices

A *v-pessimal vertex*  $v'$  is a vertex that satisfies  $H(v', v) = \max_{w \in V} H(w, v)$ . Note that pessimal vertices are not necessarily unique; for example, every leaf of the star  $S_n$  is pessimal for the central vertex. We use  $v''$  to denote a pessimal vertex for  $v'$ , and we employ the notation

$$\text{Pess}(v) = H(v', v) = \max_{w \in V} H(w, v),$$

so that we can refer to the pessimal hitting time to  $v$  in a manner agnostic of the choice of pessimal vertex  $v'$ . This allows us to define the lightweight notation

$$\Delta\text{Pess}(v) = \text{Pess}_{\tilde{G}}(v) - \text{Pess}_G(v) \tag{1}$$

which shields us from the fact that the  $v$ -pessimal vertices in  $G$  and  $\tilde{G}$  may be distinct.

### 2.2 Calculating mixing times on trees

Next we address the calculation of  $H(v, \pi)$ , the expected length of an optimal stopping rule from the singleton distribution on  $v$  to the stationary distribution  $\pi$ . We refer to  $H(v, \pi)$  as the *mixing time from v*. The following result holds for any graph  $G$ .

**Theorem 2.1 ([9])** *The expected length of an optimal mixing rule starting from vertex  $v$  is*

$$H(v, \pi) = H(v', v) - \sum_{u \in V} \pi_u H(u, v) = \text{Pess}(v) - \sum_{j \in V} \pi_j H(j, v). \tag{2}$$

The very useful formula (2) allows us to calculate the access time  $H(v, \pi)$  via a linear combination of vertex-to-vertex hitting times. In other words, if we know all the pairwise hitting times, it is easy to calculate the access time  $H(v, \pi)$  for each starting vertex  $v \in V$ .

For a given graph, we can determine every hitting time  $H(u, v)$  by solving a system of linear equations. In the special case of trees, we have an explicit formula in terms of the distances and degrees of the graph. Variations on this formula appear in the literature (cf. [3, 4, 6, 8] and Chapter 5 of [1]). This paper builds on the results in [3, 4], so we opt for that formulation. From here forward, we assume that  $G = (V, E)$  is a tree on  $n \geq 4$  vertices.

We start with a well-known result about the hitting time between adjacent vertices in trees. If  $uv \in E$ , then removing this edge breaks  $G$  into two disjoint trees  $G_u$  and  $G_v$  where  $u \in V(G_u)$  and  $v \in V(G_v)$ . We define  $V_{u:v} = V(G_u)$  and  $V_{v:u} = V(G_v)$ . We think of  $V_{u:v}$  as

the set of vertices that are closer to  $u$  than to  $v$ . For these adjacent vertices, let  $F$  denote the induced tree on  $V_{u:v} \cup \{v\}$ . We have

$$H(u, v) = \text{Ret}_F(v) - 1 = \sum_{w \in V_{u:v}} \deg(w) = 2|V_{u:v}| - 1, \quad (3)$$

where the first equality holds by the well-known equality  $\text{Ret}(v) = 2|E|/\deg(v)$ . Equation (3) encodes a very useful property of trees: the hitting time from a vertex  $u$  to an adjacent vertex  $v$  only depends on  $|V_{u:v}|$ , independent of the particular structure of the tree. This equation can be used to reveal that when  $G = P_n$  is the path  $(v_1, v_2, \dots, v_n)$ , the hitting times are

$$H_{P_n}(v_i, v_j) = \begin{cases} (i-1)^2 - (j-1)^2 & i \leq j, \\ (n-j)^2 - (n-i)^2 & i > j. \end{cases} \quad (4)$$

For example, the first of these formulas is calculated by repeatedly using equation (3) to evaluate  $H(v_i, v_j) = H(v_i, v_{i+1}) + H(v_{i+1}, v_{i+2}) + \dots + H(v_{j-1}, v_j)$ . A similar argument produces a formula for the hitting times of an arbitrary tree, but first we need some additional notation. Let  $u, v, w \in V$ . Define  $d(u, v)$  to be the distance between these two vertices and define

$$\ell(u, v; w) = \frac{1}{2}(d(u, w) + d(v, w) - d(u, v))$$

to be the length of the intersection of the  $(u, w)$ -path and the  $(v, w)$ -path.

**Lemma 2.2** ([3]) *For any pair of vertices  $v_i, v_j$ , we have*

$$H(v_i, v_j) = \sum_{v \in V} \ell(v_i, v; v_j) \deg(v). \quad (5)$$

Theorem 2.1 and Lemma 2.2 are the fundamental tools in our proof of Theorem 1.2. Together they provide a simple way to calculate state-to-state hitting times and mixing times.

### 2.3 The foci of a tree

The center and the barycenter are two well-established notions of centrality. The center of a tree  $G$  is the vertex (or two adjacent vertices) that achieves  $\min_{u \in V} \max_v d(v, u)$ . The center does not appear to have any central properties with respect to random walks on trees. The barycenter is the vertex (or two adjacent vertices) that achieves  $\min_{u \in V} \sum_{v \in V} d(v, u)$ . Proposition 1 of [3] shows that the barycenter is the “average” center for random walks on trees: this vertex also achieves  $\min_u \sum_{v \in V} \pi_v H(v, u)$ . A third type of centrality for trees was introduced in [3]: the *focus* is the “extremal” center for random walks.

**Definition 2.3** ([3]) *A vertex  $v$  achieving  $\text{Pess}(v) = \min_{w \in V} \text{Pess}(w)$  is called a primary focus of  $G$ . If all  $v$ -pessimal vertices are contained in a single subtree of  $G - v$ , then the unique  $v$ -neighbor  $u$  in that subtree is also a focus of  $G$ . If  $\text{Pess}(u) = \text{Pess}(v)$ , then  $u$  is also a primary focus. Otherwise  $u$  is a secondary focus.*

Every tree has either one focus, in which case it is *focal*, or two adjacent foci, in which case it is *bifocal*. We have the following two theorems, which relate the foci to mixing walks on  $G$ .

**Theorem 2.4 ([3])** *If  $G$  is focal with focus  $u$ , then for all  $v$   $H(v, \pi) = H(v, u) + H(u, \pi)$ . If  $G$  is bifocal with foci  $u$  and  $w$ , then for  $v \in V_{u:w}$ ,  $H(v, \pi) = H(v, u) + H(u, \pi)$  and for  $v \in V_{w:u}$ ,  $H(v, \pi) = H(v, w) + H(w, \pi)$ .*

**Theorem 2.5 ([3])** *If  $H(w, \pi) = T_{\text{bestmix}}(G)$  then  $w$  is a focus of  $G$ . For bifocal  $G$  with foci  $u$  and  $v$ , if  $H(v', u) < H(u', v)$  then  $v$  is the unique vertex achieving  $T_{\text{bestmix}}(G)$ . If  $H(v', u) > H(u', v)$  then  $u$  is the unique vertex achieving  $T_{\text{bestmix}}(G)$ . If  $H(v', u) = H(u', v)$  then both vertices achieve  $T_{\text{bestmix}}$ .*

If  $H(w, \pi) = T_{\text{bestmix}}(G)$  then we say that  $w$  is a *best mix focus*. Note that a tree can have one or two best mix foci. In the former case, the best mix focus could be either the primary focus or the secondary focus: it depends on the relative sizes of  $H(v', u)$  and  $H(u', v)$ . This unusual criterion is what makes handling the best mixing time so delicate. A small change to the tree structure can move the location of the best mix focus, even when the foci do not change.

## 2.4 The best mixing time for stars, paths and odd wishbones

As our first application of these theorems and lemmas, let us calculate the best mixing time for the star, the path (both even and odd  $n$ ) and the wishbone (for odd  $n$ ). This will justify the expressions that appear in Theorem 2.1. Clearly, the central vertex  $v$  of the star  $S_n = K_{1,n-1}$  is its unique focus. Let  $w \neq v$  be any other vertex. We have  $\pi_v = 1/2$  and  $\pi_w = 1/2(n-1)$ . Furthermore,  $H(w, v) = \text{Pess}(v) = 1$  and obviously  $H(v, v) = 0$ . Therefore

$$T_{\text{bestmix}}(S_n) = \text{Pess}(v) - \sum_{w \in V} \pi_w H(w, v) = 1 - \frac{1}{2(n-1)}(n-1) = \frac{1}{2}.$$

In fact, it is easy to see that  $S_n$  is the unique tree on  $n$  vertices that achieves this value. Let  $G \neq S_n$  and let  $v$  be a best mix focus. Then  $\pi_v = \deg(v)/2(n-1) < 1/2$  because the star is the unique tree with a vertex of degree  $n-1$ . An optimal mixing rule started at  $v$  must exit with probability at least  $1 - \pi_v > 1/2$ , otherwise the ending distribution will weight vertex  $v$  too heavily. This proves that the star is the unique minimizing structure for Theorem 1.2.

Next, we calculate  $T_{\text{bestmix}}(P_n)$ . Label the vertices as  $(v_1, v_2, \dots, v_n)$ . We first consider odd  $n = 2r + 1$ . By symmetry, the unique focus is  $v_{r+1}$ . By Theorem 2.1 and equation (4),

$$\begin{aligned} T_{\text{bestmix}}(P_{2r+1}) &= \text{Pess}(v_{r+1}) - \sum_{i=1}^{2r+1} \pi_i H(v_i, v_{r+1}) = \text{Pess}(v_{r+1}) - 2 \sum_{i=1}^r \pi_i H(v_i, v_{r+1}) \\ &= r^2 - \frac{1}{2r}r^2 - \frac{1}{r} \sum_{i=2}^r (r^2 - (i-1)^2) = \frac{2r^2 + 1}{6} = \frac{n^2 - 2n + 3}{12}, \end{aligned}$$

where the second equality follows from the symmetry of the odd path. The calculation for the even path is a bit tougher. Setting  $n = 2r$ , we have two best mix foci  $v_r, v_{r+1}$  by symmetry. We will take  $v_{r+1}$  as the best mix focus in our calculation. We have

$$\begin{aligned}
T_{\text{bestmix}}(P_{2r}) &= \text{Pess}(v_{r+1}) - \sum_{i=1}^r \pi_i H(v_i, v_{r+1}) - \sum_{i=r+2}^{2r} \pi_i H(v_i, v_{r+1}) \\
&= \text{Pess}(v_{r+1}) - 2 \sum_{i=1}^{r-1} \pi_i H(v_i, v_r) - \sum_{i=1}^r \pi_i H(v_r, v_{r+1}) \\
&= r^2 - \frac{2}{4r-2}(r-1)^2 - \frac{4}{4r-2} \sum_{i=2}^{r-1} ((r-1)^2 - (i-1)^2) - \frac{1}{2}(2r-1) \\
&= \frac{2r^2 + 4r - 3}{6} = \frac{n^2 + 4n - 6}{12}
\end{aligned}$$

where we use the symmetry of the path and  $H(v_i, v_{r+1}) = H(v_i, v_r) + H(v_r, v_{r+1})$  for  $1 \leq i < r$  in the third equality.

Finally, we determine  $T_{\text{bestmix}}(Y_{2r+1})$ . Let the spine be  $(w_1, w_2, \dots, w_{2r})$  with a leaf  $x$  adjacent to  $w_{r+1}$ . Arguing similarly to the even path, we have

$$\begin{aligned}
T_{\text{bestmix}}(Y_{2r+1}) &= \text{Pess}(w_{r+1}) - \sum_{i=1}^r \pi_i H(w_i, w_{r+1}) - \sum_{i=r+2}^{2r} H(w_i, w_{r+1}) - \pi_x H(x, w_{r+1}) \\
&= \text{Pess}(w_{r+1}) - 2 \sum_{i=1}^r \pi_i H(w_i, w_r) - \sum_{i=1}^r H(w_r, w_{r+1}) - \pi_x H(x, w_{r+1}) \\
&= r^2 - \frac{1}{2r}(r-1)^2 - \frac{1}{r} \sum_{i=2}^{r-1} ((r-1)^2 - (i-1)^2) - \frac{2r-1}{4r}(2r-1) - \frac{1}{4r} \\
&= \frac{n^2 + 4n - 15}{12}.
\end{aligned}$$

These calculations show that for odd  $n = 2r + 1 \geq 5$ , we have  $T_{\text{bestmix}}(Y_n) = T_{\text{bestmix}}(P_n) + (n-3)/2$ . Furthermore, we can identify precisely what leads to the wishbone's advantage in equation (2). The pessimal hitting times to the best mix foci of  $Y_{2r+1}$  and  $P_{2r+1}$  are both equal to  $r^2$ . So the average hitting time to the best mix focus makes the difference: this quantity is smaller for  $Y_n$ . We will see below that this effect can be attributed to the fact that  $P_{2n+1}$  has a single focus. This special balance to the tree actually makes it *easier* to mix from the focus.

With these calculations in hand, we have proven the lower bound of Theorem 1.2, and calculated the two expressions that appear in the upper bound. The remainder of the paper proves that these upper bounds are correct.

### 3 Caterpillars and Tree Surgeries

In this section, we show that for a fixed  $n \geq 4$ , the tree on  $n$  vertices that maximizes the best mixing time is a caterpillar. Next, we describe our basic techniques, called *tree surgeries*, for

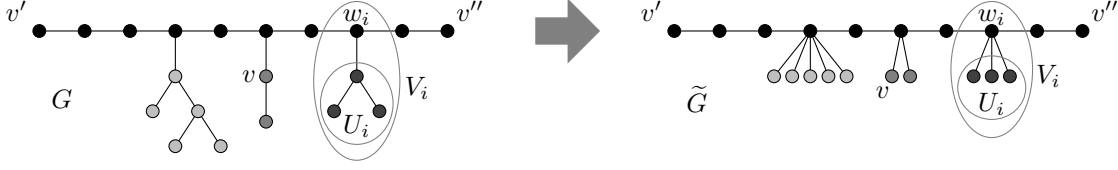


Figure 3.1: Transforming a tree  $G$  into a caterpillar  $\tilde{G}$  with  $T_{\text{bestmix}}(G) < T_{\text{bestmix}}(\tilde{G})$ . Every vertex in  $U_i$  becomes adjacent to  $w_i$ .

making incremental changes to a caterpillars. We introduce some notation and useful lemmas for showing that the best mixing time monotonically increases after each tree surgery.

### 3.1 Tree to Caterpillar

We begin with a lemma showing that  $T_{\text{bestmix}}$  is maximized by a caterpillar. Recall that a *caterpillar* is a tree such that every vertex is distance at most one from a fixed central path  $W = \{w_1, w_2, \dots, w_t\}$ , called the *spine*. We employ the following notation for talking about sections of the caterpillar. For  $2 \leq i \leq t-1$ , let  $U_i \subset V \setminus W$  denote the set of pendant leaves adjacent to  $w_i$ . Note that the spinal leaves are not in  $V \setminus W$ , so  $w_1 \notin U_2$  and  $w_t \notin U_{t-1}$ . Finally, we also define  $V_k = U_k \cup \{w_k\}$ . See Figure 3.1 for an example of  $w_i$  and  $U_i \subset V_i$ .

**Lemma 3.1** *Let  $G$  be a tree on  $n$  vertices that is not a caterpillar. There exists a caterpillar  $\tilde{G}$  on  $n$  vertices such that  $T_{\text{bestmix}}(\tilde{G}) > T_{\text{bestmix}}(G)$ .*

**Proof:** We give a simple construction of the caterpillar  $\tilde{G}$ ; an example is shown in Figure 3.1.

Choose any  $v \in V(G)$ . Recall from Section 2.1 that  $v'$  is a  $v$ -pessimal vertex, and that  $v''$  is a  $v'$ -pessimal vertex. Let  $W = \{v' = w_1, w_2, \dots, w_t = v''\}$  be the vertices on the unique  $(v', v'')$  path  $P \subset G$ . These vertices will be the spine of the caterpillar  $\tilde{G}$ . (Note that  $v$  is not necessarily a member of  $W$ .) We replace the subforest attached to each spine vertex  $w_i$  with a set of leaves adjacent to  $w_i$ . Define  $\tilde{G}$  to be the caterpillar with vertex set  $V(G)$  and edge set  $E(\tilde{G}) = E(P) \cup \{uw_i \mid u \in U_i, 1 < i < t\}$ .

The foci of  $\tilde{G}$  are contained in  $W$  because leaves cannot be foci when  $n \geq 3$ . Without loss of generality, let  $w_r$  be the best mix focus of  $\tilde{G}$  and let  $w'_r = w_1$ . If  $G$  is bifocal, the other focus must be  $w_{r-1}$  and  $w'_{r-1} = w_t$ . For  $1 \leq i \leq t-1$ , the quantity

$$\sum_{v \in V_{w_i:w_{i+1}}} \deg_G(v) = 2|V_{w_i:w_{i+1}}| - 1 = \sum_{v \in V_{w_i:w_{i+1}}} \deg_{\tilde{G}}(v)$$

does not change. By equation (3) we have

$$H_G(w_i, w_j) = H_{\tilde{G}}(w_i, w_j) \tag{6}$$

for all  $1 \leq i, j \leq t$ . Considering any non-spine vertex  $v \in U_i$ , we have

$$H_{\tilde{G}}(v, w_r) - H_G(v, w_r) = 1 + H_{\tilde{G}}(w_i, w_r) - (H_G(v, w_i) + H_G(w_i, w_r)) = 1 - H_G(v, w_i) \leq 0.$$



We conclude that  $w_1$  is  $v_r$ -pessimal for both  $G$  and  $\tilde{G}$ . Likewise,  $w_t$  is  $w_{r-1}$ -pessimal for both  $G$  and  $\tilde{G}$ . By equation (6) and Theorem 2.5, the vertex  $w_r$  is the best mix focus of  $\tilde{G}$  and  $\text{Pess}_G(w_r) = \text{Pess}_{\tilde{G}}(w_r)$ . Finally, Theorem 2.1 yields

$$\begin{aligned} T_{\text{bestmix}}(\tilde{G}) - T_{\text{bestmix}}(G) &= H_{\tilde{G}}(w_r, \tilde{\pi}) - H_G(w_r, \pi) \\ &= \sum_{i=1}^t \sum_{v \in V_i} (\pi_v(H_G(v, w_i) + H_G(w_i, w_r)) - \tilde{\pi}_v(H_{\tilde{G}}(v, w_i) + H_{\tilde{G}}(w_i, w_r))) \\ &= \sum_{i=1}^t \sum_{v \in V_i} (\pi_v H_G(v, w_i) - \tilde{\pi}_v H_{\tilde{G}}(v, w_i)) = \sum_{i=2}^{t-1} \sum_{v \in U_i} (\pi_v H_G(v, w_i) - \tilde{\pi}_v H_{\tilde{G}}(v, w_i)) \geq 0 \end{aligned}$$

because each  $\sum_{v \in V_i} \pi_v H_G(v, w_i) - \tilde{\pi}_v H_{\tilde{G}}(v, w_i) \geq 0$ , with equality holding if and only if every  $v \in U_i$  is already adjacent to  $w_i$ . Since  $G$  was not a caterpillar, there must be at least one exceptional  $v$  in some  $U_i$ , so this inequality is strict.  $\blacksquare$

### 3.2 Leaf Transplants and Tree Surgeries

For the remainder of the paper,  $G$  is a caterpillar on  $n$  vertices with spine  $W = \{w_1, w_2, \dots, w_t\}$ . We choose to view  $w_1$  as the leftmost vertex and  $w_t$  as the rightmost vertex. For  $v \in V$ , we define the *left pessimal hitting time*  $L(v) = H(w_1, v)$  and the *right pessimal hitting time*  $R(v) = H(w_t, v)$ . Of course,  $\text{Pess}(v) = \max\{L(v), R(v)\}$ . When  $G$  is bifocal, we always label the foci as  $w_{r-1}$  and  $w_r$  where  $w_r$  achieves  $H(w_r, \pi) = T_{\text{bestmix}}(G)$ . In this case, we view  $\{w_1, \dots, w_{r-2}\}$  as the *left spine* and  $\{w_{r+1}, \dots, w_t\}$  as the *right spine*. The foci  $\{w_{r-1}, w_r\}$  are considered the *central spine*. Note that this arrangement guarantees that  $w'_r = w_1$  and  $w'_{r-1} = w_t$ . We collect some basic results about caterpillar foci in the next lemma.

**Lemma 3.2** *Let  $G$  be a caterpillar with spine  $W = \{w_1, w_2, \dots, w_t\}$ . Then*

- (a) *The vertex  $w_r$  is the unique focus of  $G$  if and only if  $L(w_r) = R(w_r)$ .*
- (b) *The foci of  $G$  are  $w_{r-1}, w_r$  if and only if  $L(w_r) > R(w_r)$  and  $R(w_{r-1}) > L(w_{r-1})$ .*
- (c) *If  $G$  has foci  $w_{r-1}, w_r$  then vertex  $w_r$  is a best mix focus if and only if  $L(w_{r-1}) \leq R(w_r)$ .*
- (d) *If  $L(w_r) > R(w_r)$  and  $L(w_{r-1}) \leq R(w_r)$  then  $w_r$  is a best mix focus and  $w_{r-1}$  is also a focus of  $G$ .*

**Proof:** Parts (a) and (b) are a direct consequence of Definition 2.3. Part (c) is restatement of Theorem 2.5 for caterpillars. For part (d), we observe that  $L(w_{r-1}) \leq R(w_r) < R(w_{r-1})$ , so that  $w_{r-1}$  and  $w_r$  satisfy (b) and (c).  $\blacksquare$

When  $U_i \neq \emptyset$ , we define the *leaf transplant*  $\tau(i, j)$  as the relocation of  $x \in U_i$  so that this leaf is now adjacent to  $w_j$  where  $1 \leq j \leq t$ . Usually, we have  $2 \leq j \leq t-1$ , so that the leaf  $x$  becomes an element of  $U_j$ . Our most common operation will be an *elementary leaf transplant*



Figure 3.2: The effect of transplants on hitting times. The right tree results from applying  $\tau(2, 5) \wedge \tau(8, 9)$  on the left tree. Each spine edge  $(w_i, w_{i+1})$  is labeled by  $H(w_i, w_{i+1})$  above and  $H(w_{i+1}, w_i)$  below. Other hitting times are linear combinations of those shown.

where  $j = i \pm 1$ . Occasionally, we will take  $j \in \{1, t\}$ , so that the leaf transplant actually increases the length of the spine. Finally, we need one special transplant that reduces the length of the spine by relocating the leaf  $w_t$  to become adjacent to the best mix focus  $w_r$ : we use  $\sigma(t, r)$  to denote this rare spinal leaf transplant. The more general term *tree surgery* denotes either a single leaf transplant (either  $\tau(i, j)$  or  $\sigma(t, r)$ ), or a pair of leaf transplants  $\tau(i_1, j_1) \wedge \tau(i_2, j_2)$  performed simultaneously. We use  $\mathcal{S}(G)$  to denote the caterpillar that results from applying tree surgery  $\mathcal{S}$  to caterpillar  $G$ . Our algorithm uses eleven different tree surgeries. They are enumerated in Tables 4.1, 4.2 and 4.3 below.

The remainder of this section is devoted to proving some fundamental results about the effects of tree surgeries on the best mixing time. Let  $G = (V, E)$  be a caterpillar and let  $\tilde{G} = \mathcal{S}(G) = (V, \tilde{E})$  be the caterpillar obtained by applying tree surgery  $\mathcal{S}$ , with stationary distribution  $\tilde{\pi} = \pi_{\tilde{G}}$ . We introduce notation for the comparison of various caterpillar measurements. Just as with the pessimal notation introduced in Section 2.1, we are motivated to shield ourselves from the particular locations of important vertices (like foci and pessimal vertices) which may be different in  $G$  and  $\tilde{G}$ .

As with equation (1), we introduce  $\Delta$ -notation to compactly represent the changes in various quantities due to a tree surgery. For vertices  $v, v_1, v_2 \in V$ , we define:

$$\begin{aligned} \Delta \deg(v) &= \deg_{\tilde{G}}(v) - \deg_G(v), & \Delta \text{Pess}(v) &= \text{Pess}_{\tilde{G}}(v) - \text{Pess}_G(v), \\ \Delta \pi(v) &= \tilde{\pi}(v) - \pi(v), & \Delta L(v) &= L_{\tilde{G}}(v) - L_G(v), \\ \Delta H(v_1, v_2) &= H_{\tilde{G}}(v_1, v_2) - H_G(v_1, v_2), & \Delta R(v) &= R_{\tilde{G}}(v) - R_G(v), \\ \Delta H(\pi, v) &= H_{\tilde{G}}(\tilde{\pi}, v) - H_G(\pi, v), & \Delta T_{\text{bestmix}} &= T_{\text{bestmix}}(\tilde{G}) - T_{\text{bestmix}}(G). \end{aligned}$$

Let  $G$  be a caterpillar and let  $\tilde{G} = \mathcal{S}(G)$  be the result of a leaf transplant  $\tau(i, j)$ . We give formulas for  $\Delta H(v, w_r)$  for use in later sections. We only consider leaf transplants  $\tau(i, j)$  where either  $1 \leq i, j \leq r$  or where  $r+1 \leq i, j \leq t$  (so that we never transplant leaves from the left spine to the right spine, or vice versa). The formulas below cover two qualitatively different cases for  $\tau(i, j)$ . When  $2 \leq j \leq t-1$ , the spine length is unaffected. However, when  $j \in \{1, t\}$ , the spine length increases. The formulas below still hold, though the arguments are slightly different. All four formulas below follow quickly from equation (3), and we leave these short calculations to the reader. Figure 3.2 shows the effect of two example transplants on hitting times. The arguments for  $j = i \pm 1$  are straight forward, and the other cases follow inductively.

Suppose that the leaf transplant  $\tau(i, j)$  moves  $x$  from  $U_i$  to  $U_j$ . For  $v \in V_k$ , the value of  $\Delta H(v, w_r)$  depends on the relative locations of  $i, j$  and  $k$ :

$$\text{If } 1 \leq j < i \leq r \text{ then } \Delta H(v, w_r) = \begin{cases} H_G(w_j, w_i) + 2(i - j) & v = x, \\ 2(i - j) & k \leq j, \\ 2(i - k) & j < k < i, \\ 0 & k \leq i \leq r \text{ and } v \neq x. \end{cases} \quad (7)$$

$$\text{If } 1 \leq i < j \leq r \text{ then } \Delta H(v, w_r) = \begin{cases} -H_G(w_i, w_j) & v = x, \\ -2(j - i) & k \leq i \text{ and } v \neq x, \\ -2(j - k) & i < k < j, \\ 0 & j \leq k \leq r. \end{cases} \quad (8)$$

$$\text{If } r \leq j < i \leq t \text{ then } \Delta H(v, w_r) = \begin{cases} -H_G(w_i, w_j) & v = x, \\ -2(i - j) & i \leq k \text{ and } v \neq x, \\ -2(i - k) & j < k < i, \\ 0 & r < k \leq j. \end{cases} \quad (9)$$

$$\text{If } r \leq i < j \leq t \text{ then } \Delta H(v, w_r) = \begin{cases} H_G(w_j, w_i) + 2 & v = x, \\ 2(j - i) & j \leq k, \\ 2(j - k) & i < k < j, \\ 0 & r < k \leq i \text{ and } v \neq x. \end{cases} \quad (10)$$

We get similar equations for hitting times to the second focus  $w_{r-1}$ , with the appropriate change to the bounds on the indices  $i, j$ . This completes our list of useful hitting time changes due to common leaf transplants.

Next, we turn our attention to tracking the change in the best mixing time after a tree surgery. The lemmas that follow will be used to analyze all eleven surgeries used in our algorithm. If  $w_r$  is the best mix focus of  $G$  and  $w_s$  is the best mix focus of  $\tilde{G}$ , then

$$\Delta T_{\text{bestmix}} = \text{Pess}_{\tilde{G}}(w_s) - \text{Pess}_G(w_r) - (H_{\tilde{G}}(\tilde{\pi}, w_s) - H_G(\pi, w_r)). \quad (11)$$

As we alter our caterpillar, we strive to maintain  $w_r$  as the best mix focus and the leftmost spinal vertex as the  $w_r$ -pessimal vertex. In this case, equation (11) simplifies to  $\Delta T_{\text{bestmix}} = \Delta \text{Pess}(w_r) - \Delta H(\pi, w_r) = \Delta L(w_r) - \Delta H(\pi, w_r)$ . This can be rephrased as the simple criterion:

$$\Delta T_{\text{bestmix}} \geq 0 \iff \Delta \text{Pess}(w_r) \geq \Delta H(\pi, w_r). \quad (12)$$

The next two results track the effect of a surgery on the location of the foci.

**Lemma 3.3** *Let  $G$  be a caterpillar. Let  $\mathcal{S}$  be a tree surgery and let  $\tilde{G} = \mathcal{S}(G)$ . If*

$$L_G(w_r) - R_G(w_r) > \Delta R(w_r) - \Delta L(w_r) \quad (13)$$

$$\text{and } R_G(w_r) - L_G(w_{r-1}) \geq \Delta L(w_{r-1}) - \Delta R(w_r), \quad (14)$$

*then  $\tilde{G} = \mathcal{S}(G)$  has best mix focus  $w_r$  and  $w_{r-1}$  is also a focus.*

**Proof:** These criteria are equivalent to the conditions of Lemma 3.2 (d) for  $\tilde{G}$ . ■

**Corollary 3.4** *Let  $G$  be a bifocal caterpillar with focus  $w_{r-1}$  and best mix focus  $w_r$ . Let  $\mathcal{S}$  be a tree surgery and let  $\tilde{G} = \mathcal{S}(G)$ . If  $\Delta R(w_r) - \Delta L(w_r) \leq 0$  and  $\Delta L(w_{r-1}) - \Delta R(w_r) \leq 0$  then  $\tilde{G} = \mathcal{S}(G)$  also has focus  $w_{r-1}$  and best mix focus  $w_r$ .*

**Proof:** This follows directly from Lemma 3.3 and the inequalities  $L(w_r) - R(w_r) \geq 1$  and  $R(w_r) - L(w_{r-1}) \geq 0$ .  $\blacksquare$

Lemma 3.3 and Corollary 3.4 are key tools of our methodology. We use them repeatedly to verify the foci and best mix focus of our caterpillar after a surgery has been applied. Once we know that  $w_r$  is still the best mix focus, we then verify that equation (12) holds. Next, we give a final test for  $\Delta T_{\text{bestmix}} \geq 0$ . This lemma will be used very frequently in subsequent sections.

**Lemma 3.5** *Let  $G$  be a caterpillar and let  $\mathcal{S}$  be a tree surgery such that  $w_r$  is a best mix focus of  $G$  and  $\tilde{G}$ . Let  $A = \{v \in V \mid \Delta H(v, w_r) > \Delta \text{Pess}(w_r)\}$  and let  $C = \{v \in \bar{A} \mid \Delta(v) > 0\}$ . If there exists  $C \subset B \subset \bar{A}$  for which*

$$\begin{aligned} & \sum_{v \in A} \left( \deg_{\tilde{G}}(v)(\Delta H(v, w_r) - \Delta \text{Pess}(w_r)) + \Delta \deg_G(v)H_G(v, w_r) \right) \\ & \leq \sum_{u \in B} \left( \deg_{\tilde{G}}(u)(\Delta \text{Pess}(w_r) - \Delta H(u, w_r)) - \Delta \deg_G(u)H_G(u, w_r) \right) \end{aligned} \quad (15)$$

*then  $\Delta T_{\text{bestmix}} \geq 0$ .*

Before proving this lemma, we make a few comments its use in later sections. First, the set  $A$  will always be small: it will only contain one or both of the leaves moved during the surgery. Second, if the lemma holds for  $B$ , then it holds for any superset of  $B$ , including  $\bar{A}$ . However, there is no need to calculate the right hand side for  $\bar{A}$  when a small subset will do. For our earlier surgeries, the set  $B$  will be small, typically consisting of a handful of spine vertices near the transplant locations. In later surgeries,  $B$  will consist of half or all of the spine.

**Proof:** We decompose  $\Delta H(\pi, w_r)$  as follows:

$$\begin{aligned} \Delta H(\pi, w_r) &= \sum_{v \in V} (\tilde{\pi}(v)H_{\tilde{G}}(v, w_r) - \pi(v)H_G(v, w_r)) \\ &= \sum_{v \in V} (\tilde{\pi}(v)(\Delta H(v, w_r) + H_G(v, w_r)) - \pi(v)H_G(v, w_r)) \\ &= \sum_{v \in V} (\tilde{\pi}(v)\Delta H(v, w_r) + \Delta \pi(v)H_G(v, w_r)). \end{aligned}$$

Let  $g(v) = \deg_{\tilde{G}}(v)(\Delta \text{Pess}(w_r) - \Delta H(v, w_r)) - \Delta \deg_G(v)H_G(v, w_r)$ . We can rewrite

$$\begin{aligned} \Delta \text{Pess}(w_r) - \Delta H(\pi, w_r) &= \Delta \text{Pess}(w_r) - \sum_{v \in V} (\tilde{\pi}(v)\Delta H(v, w_r) + \Delta \pi(v)H_G(v, w_r)) \\ &= \frac{1}{2|E|} \left( \sum_{v \in A} g(v) + \sum_{u \in \bar{A}} g(u) \right) \geq \frac{1}{2|E|} \left( \sum_{v \in A} g(v) + \sum_{u \in B} g(u) \right). \end{aligned}$$

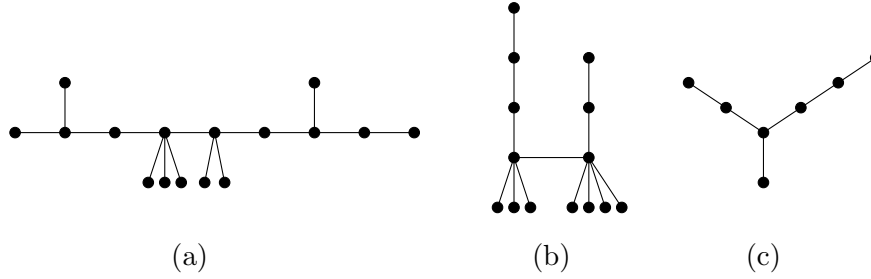


Figure 4.1: Examples of the milestone graphs obtained at the end of each phase of our process: (a) a seesaw, (b) a twin broom, and (c) the wishbone  $Y_7$ .

The final inequality holds because  $g(u) \geq 0$  for every  $u \in \overline{A} \setminus B \subset \overline{A} \setminus C$ . Therefore, if  $-\sum_{v \in A} g(v) \leq \sum_{u \in B} g(u)$  then  $\Delta \text{Pess}(w_r) - \Delta H(\pi, w_r) \geq 0$ . ■

We have an immediate corollary for the special case where  $A$  is empty.

**Corollary 3.6** *Given a tree  $G$  and any tree surgery  $\mathcal{S}$  such that  $w_r$  is the best mix focus of  $G$  and  $\tilde{G}$ , if  $\Delta H(v, w_r) \leq \Delta \text{Pess}(w_r)$ , for all  $v \in V$ , then  $\Delta T_{\text{bestmix}} \geq 0$ .* □

Finally, we note that if the  $w_r$ -pessimal vertices in  $G$  and  $\tilde{G}$  are the leftmost spinal vertices in these graphs, then we can replace  $\Delta \text{Pess}(w_r)$  with  $\Delta L(w_r)$  in Lemma 3.5 and Corollary 3.6.

## 4 Proof of the upper bound in Theorem 1.2

In Section 2.4, we proved the lower bound of Theorem 1.2, and calculated the best mix time for the even path and the odd wishbone. In this section, we prove the upper bound of Theorem 1.2, leaving the details to the lemmas in the accompanying subsections. Two examples of the algorithm in practice are shown in Figure 1.1 above.

**Proof of Theorem 1.2:** Given a tree  $G$  that is not an even path or an odd wishbone, we must show that it does not maximize the best mixing time. We may assume that  $G$  is not a star since that is clearly not the maximizing structure. Starting from our tree  $G$ , we let  $G_0$  be the caterpillar constructed by Lemma 3.1, so that  $T_{\text{bestmix}}(G) \leq T_{\text{bestmix}}(G_0)$ . Next, we apply a sequence of tree surgeries to produce a sequence of caterpillars  $G_0, G_1, G_2, \dots, G_m$  such that  $T_{\text{bestmix}}(G_{i-1}) \leq T_{\text{bestmix}}(G_i)$ , where the final caterpillar  $G_m$  is either  $P_n$  or  $Y_n$ . This occurs in an algorithmic manner, divided into three phases.

In order to define these phases, we define two special subfamilies of caterpillars which mark milestones in our graph sequence. Examples of these graphs and a wishbone are shown in Figure 4.1. A *twin broom* is a bifocal caterpillar whose only non-spine leaves are adjacent to the foci  $w_{r-1}, w_r$ . In other words,  $V = W \cup U_{r-1} \cup U_r$ . A *seesaw* is a twin broom that has at most one additional leaf on each side of the spine. In other words, a seesaw satisfies  $|\cup_{i=2}^{r-2} U_i| \leq 1$  and  $|\cup_{i=r+1}^{t-1} U_i| \leq 1$ . Figure 1.1 includes examples of each. In that figure, the

output of  $S_4$  and the initial tree on 11 vertices are both seesaws. Meanwhile, the outputs of  $S_6$  and  $S_8$  are both twin brooms (as are all graphs that follow them).

We start with a caterpillar  $G$  on  $n$  vertices. In Phase One, we convert the caterpillar into a seesaw. Phase Two converts the seesaw into a twin broom. Phase Three converts the twin broom into one of  $P_n$  (when  $n$  is even) or  $Y_n$  (when  $n$  is odd). Lemmas 4.1, 4.7, and 4.11 below show that the best mixing time monotonically increases in every phase of this process. This proves Theorem 1.2.  $\blacksquare$

Our caterpillar transformation consists of incremental steps that move one or two leaves at a time. This allows us to monitor the delicate balance maintained by the best mix focus. In particular, there may be a critical step at which we change the focus that attains the best mixing time. This happens in one of two ways. Usually, we make a small change that keeps the current best mix focus, but also causes a neighbor to also become a best mix focus. The other change is more abrupt: Surgery  $S_5$  below changes the location of the unique best mix focus to a neighbor of the current one. In  $S_5$ , the prescribed caterpillar structure during that surgery makes this crucial transition manageable.

The remainder of this paper is devoted to proving Lemmas 4.1, 4.7, and 4.11.

#### 4.1 Phase One: Caterpillar to Seesaw

In this subsection, we prove that Phase One is successful: we can transform any caterpillar into a seesaw while also increasing the best mixing time. Let  $G$  be a caterpillar with spine  $W = \{w_1, w_2, \dots, w_t\}$  where  $w_r$  is the best mix focus and  $w'_r = w_1$ .

**Lemma 4.1** *Let  $G$  be a caterpillar  $G$  on  $n$  vertices. Phase One creates a seesaw  $\tilde{G}$  such that  $T_{\text{bestmix}}(\tilde{G}) \geq T_{\text{bestmix}}(G)$ .*

**Proof:** If  $G$  is already a seesaw then  $\tilde{G} = G$ . Table 4.1 shows the five surgery types employed during Phase One. We defer the proofs that  $\Delta T_{\text{bestmix}} \geq 0$  for each of these surgeries to the lemmas that follow.

Figure 4.2 shows the workflow for Phase One. First, if the caterpillar has a unique focus then we use  $S_1$  to create a caterpillar with two foci. From here forward, the caterpillar will remain bifocal. Let  $w_{r-1}, w_r$  be the foci of a bifocal caterpillar with  $w_r$  achieving  $T_{\text{bestmix}}$ . A leaf  $x \in V \setminus W$  is *good* when  $x \in U_2 \cup U_{r-1} \cup U_r \cup U_{t-1}$ . All other leaves in  $V \setminus W$  are *bad*. First, we repeatedly use  $S_2$  to move pairs of bad leaves on the same side of the spine (one towards the end and the other towards the center). This loop terminates whether there is at most one bad leaf on each side of the spine. At this point, we use  $S_3$  to extend the left spine and to transplant a leaf to  $U_r$ . This requires there are at least two left leaves (which can only happen when at least one is in  $U_2$  since we are done with  $S_2$ ). After extending the spine, the previously good leaves at  $U_2$  become bad. This throws us back into the  $S_2$  loop.

We exit the  $S_3$  loop when there is at most one left leaf. At this point, we deal with the right leaves. When  $L(w_r) > R(w_r) + 1$ , we apply  $S_4$ , the right-hand surgery analogous to  $S_3$ .

However, if  $L(w_r) = R(w_r) + 1$  then applying  $S_4$  would create a focal caterpillar, which we choose to avoid. Instead, we apply  $S_5$ , which transplants a single right leaf to the end of the right spine. Applying either  $S_4$  or  $S_5$  might create bad right leaves, which puts us back into the  $S_2$  loop. Surgeries  $S_3, S_4, S_5$  reduce the number of non-spinal vertices, so Phase One must terminate. Ultimately, we create a caterpillar with at most one left leaf and at most one right leaf, while  $U_{r-1} \cup U_r$  may contain many leaves. This is a seesaw graph, as desired. ■

Surgery	Initial Conditions	Illustration
$\mathcal{S}_1$	$G$ is focal, so $L(w_r) = R(w_r)$ . The transplant depends on whether $U_{t-1} = \emptyset$ .	
$\mathcal{S}_2$	$G$ has two (or more) bad left leaves or has two (or more) bad right leaves; $R(w_r) \geq L(w_{r-1})$ .	
$\mathcal{S}_3$	$G$ has $x \in U_2$ and another left leaf $y$ ; $R(w_r) > L(w_{r-1})$	
$\mathcal{S}_4$	$G$ has $y \in U_{t-1}$ and another right leaf $x$ ; $R(w_r) \geq L(w_{r-1})$ and $L(w_r) > R(w_r) + 1$	
$\mathcal{S}_5$	$G$ has $y \in U_{t-1}$ and another right leaf $x$ ; $R(w_r) \geq L(w_{r-1})$ and $L(w_r) = R(w_r) + 1$	

Table 4.1: The Phase One tree surgeries. Except for  $\mathcal{S}_1$ , the caterpillar is bifocal with best mix focus  $w_r$  and another focus  $w_{r-1}$ . White vertices are foci and the circled vertices are best mix foci.

Next, we prove that surgeries  $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4, \mathcal{S}_5$  each result in  $\Delta T_{\text{bestmix}} \geq 0$ . If we start with a focal caterpillar  $G$ , we use surgery  $\mathcal{S}_1$  to create a bifocal caterpillar  $\tilde{G} = \mathcal{S}_1(G)$ . Depending on the structure of  $G$ , we use one of two leaf transplants:  $\tau(t-1, r)$  or  $\sigma(t, r)$ . We note that  $\sigma(t, r)$  is the only transplant that removes a leaf from the spine.

**Lemma 4.2** *Let  $G$  be a caterpillar with a single focus  $w_r$  and with the spine indexed such that  $d(w_1, w_r) \geq d(w_t, w_r)$ . If  $U_{t-1} \neq \emptyset$ , then let  $\mathcal{S}_1 = \tau(t-1, r)$ . If  $U_{t-1} = \emptyset$  then let  $\mathcal{S}_1 = \sigma(t, r)$ . The caterpillar  $\tilde{G} = \mathcal{S}_1(G)$  is bifocal and  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** We use Lemma 3.3 for our proof. The spine is indexed so that  $r-1 \geq t-r$  and the unique focus means that  $L_G(w_r) = R_G(w_r)$ . For both surgeries under consideration,  $\Delta R(w_r) < 0$  and  $\Delta L(w_r) = 0$ , so equation (13) is satisfied. Next, we observe that  $R(w_r) - L(w_{r-1}) = L(w_r) - L(w_{r-1}) = H(w_{r-1}, w_r) = \sum_{v \in V_{w_{r-1}:w_r}} \deg(v) \geq 2(r-1) - 1$  by equation (3), so it suffices to show that  $\Delta L(w_{r-1}) - \Delta R(w_r) = -\Delta R(w_r) \leq 2(r-1) - 1$  to verify equation (14). There are two cases. First, suppose that  $U_{t-1} \neq \emptyset$ , so that  $\mathcal{S}_1 = \tau(t-1, r)$  does not alter the

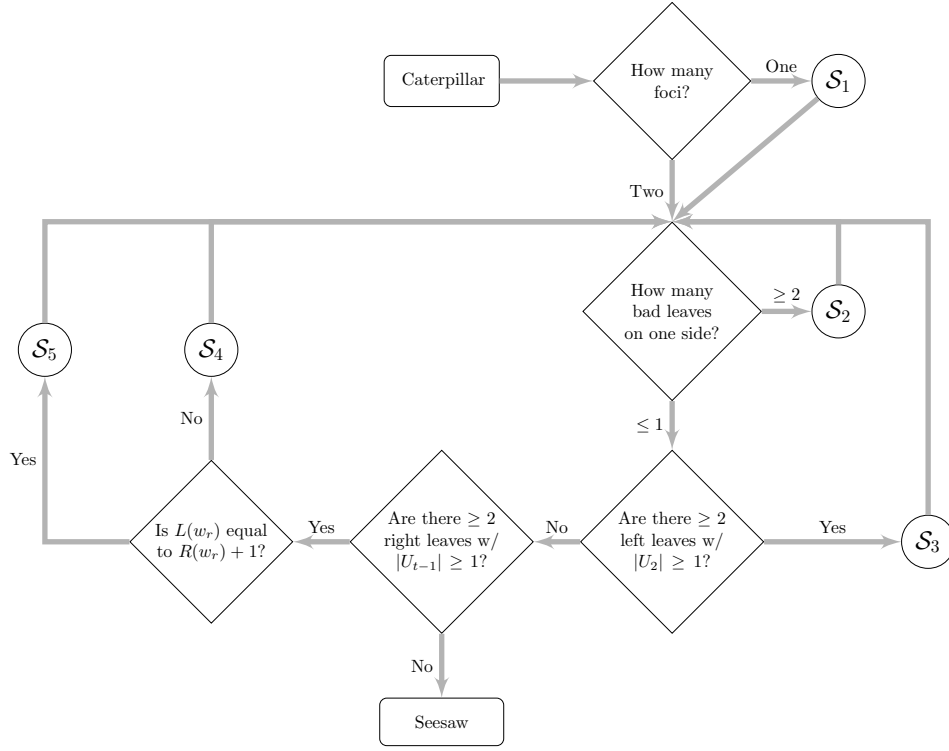


Figure 4.2: Phase One of the algorithm, turning a caterpillar into a seesaw.

spine. By equation (10), we have  $-\Delta R(w_r) = 2(t - 1 - r) < 2(r - 1) - 1$ . Next, suppose that  $U_{t-1} = \emptyset$ , so that  $\mathcal{S}_1 = \sigma(t, r)$  does alter the spine. We have

$$\begin{aligned}
 -\Delta R(w_r) &= -H_G(w_t, w_r) + H_{\tilde{G}}(w_{t-1}, w_r) = -1 - \Delta H(w_{t-1}, w_r) \\
 &= -1 + 2(t - 1 - r) = 2(t - r) - 3 < 2(r - 1) - 1.
 \end{aligned}$$

In either case, the conditions of Lemma 3.3 are satisfied, so the foci do not change. The result now follows from Corollary 3.6 since  $\Delta L(w_r) = 0 \geq \Delta H(v, w_r)$  for all  $v \in V$ .  $\blacksquare$

Next, we discuss surgery  $\mathcal{S}_2$  which transplants a pair of bad left-hand leaves, moving one toward  $U_2$  and one toward  $U_{r-1}$ , as shown in Table 4.2. We repeat  $\mathcal{S}_2$  until the left side of the caterpillar has at most one bad leaf.

**Lemma 4.3** *Let  $G$  be a bifocal caterpillar with leaves  $x \in U_i$  and  $y \in U_j$  where either  $2 < i \leq j < r - 1$  or  $r < i \leq j < t - 1$ . If  $\mathcal{S}_2 = \tau(i, i - 1) \wedge \tau(j, j + 1)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** We consider the left-hand spine case  $2 < i \leq j < r - 1$ . The right-hand spine proof is analogous, switching the roles of  $i$  and  $j$ . Suppose that there are bad leaves  $x \in U_i$  and  $y \in U_j$ . First, we observe that  $\Delta L(w_r) = 0 = \Delta L(w_{r-1})$ : these net hitting time changes are  $2 - 2 = 0$  as per equations (7) and (8). The right hand spine is unaffected, so  $\Delta R(w_r) = 0 = \Delta R(w_{r-1})$ . Corollary 3.4 guarantees that  $\tilde{G}$  also has foci  $w_{r-1}$  and best mix focus  $w_r$ .



Having established that the foci do not change, we show that  $T_{\text{bestmix}}$  increases. We must show that  $\Delta L(w_r) = \Delta H(w_1, w_r) > \Delta H(\pi, w_r)$ . We only argue the case  $i < j$ , as the case  $i = j$  is a straight-forward adaptation. Equations (7) and (8) give

$$\begin{aligned} \Delta H(x, w_r) &= H_G(w_{i-1}, w_i), & \Delta H(v, w_r) &= -2 \quad \text{for } v \in (V_i \cup \dots \cup V_j) \setminus \{x, y\}, \\ \Delta H(y, v_r) &= 2 - H_G(w_j, w_{j+1}), & \Delta H(v, w_r) &= 0 \quad \text{for all other } v. \end{aligned}$$

The only degree changes are  $\Delta \deg(w_i) = -1 = \Delta \deg(w_j)$  and  $\Delta \deg(w_{i-1}) = 1 = \Delta \deg(w_{j+1})$ . (When  $i = j$ , the only negative degree change is  $\Delta \deg(w_i) = -2$ .) We use Lemma 3.5 with  $A = \{x\}$  to show that  $T_{\text{bestmix}}$  does not decrease. The left side of inequality (15) simplifies to  $\Delta H(x, w_r) = H_G(w_{i-1}, w_i)$ . It remains to show that this value is a lower bound for the right side of (15). Taking  $B = \{w_{i-1}, w_i, w_j, w_{j+1}, y\}$ , we obtain

$$\begin{aligned} & -H_G(w_{i-1}, w_r) + (2 \deg_{\tilde{G}}(w_i) + H_G(w_i, w_r)) + (2 \deg_{\tilde{G}}(w_j) + H_G(w_j, w_r)) \\ & \quad - H_G(w_{j+1}, w_r) - (2 - H_G(w_j, w_{j+1})) \\ & = 2(\deg_G(w_i) + \deg_G(w_j) - 2) - 2 + H_G(w_i, w_r) - H_G(w_{i-1}, w_r) \\ & \quad + H_G(w_j, w_r) - H_G(w_{j+1}, w_r) + H_G(w_j, w_{j+1}) \\ & > 2H_G(w_j, w_{j+1}) - H_G(w_{i-1}, w_i) > H_G(w_{i-1}, w_i), \end{aligned}$$

where the last inequality uses equation (3) twice to justify  $H_G(w_{i-1}, w_i) < H_G(w_j, w_{j+1})$ . Thus the condition of Lemma 3.5 is satisfied, so  $\Delta T_{\text{bestmix}} > 0$ .  $\blacksquare$

Once there is at most one bad left leaf, we increase the spine length, starting with the left side. Surgery  $\mathcal{S}_3$  requires at least two left leaves, one of which must be in  $U_2$ . We also require  $R(w_r) > L(w_{r-1})$ , meaning that  $w_r$  is the unique best mix focus. If this is not the case (that is,  $R(w_r) = L(w_{r-1})$ ), then we will take  $w_{r-1}$  to be the best mix focus, reverse the labeling of the spine, and then apply  $\mathcal{S}_4$  below.

**Lemma 4.4** *Let  $G$  be a bifocal caterpillar with  $R(w_r) > L(w_{r-1})$  and with distinct vertices  $x \in U_2$  and  $y \in U_i$  where  $2 \leq i \leq r-2$ . If  $\mathcal{S}_3 = \tau(2, 1) \wedge \tau(i, i+1)$ , then  $\Delta T_{\text{bestmix}} \geq 0$ .*

**Proof:** First, we use Lemma 3.3 to show that the foci do not change. This surgery extends the left-hand side of the spine. The left-pessimal hitting times to  $w_{r-1}$  and  $w_r$  increase by  $\Delta L(w_r) = \Delta L(w_{r-1}) = 3 - 2 = 1$  (using equation (3) for the effect of  $\tau(2, 1)$  and equation (7) for  $\tau(i, i+1)$ ). The right-pessimal hitting times do not change,  $\Delta R(w_r) = \Delta R(w_{r-1}) = 0$ . Therefore, equation (13) is satisfied. By assumption,  $R(w_r) - L(w_{r-1}) \geq 1 = \Delta L(w_{r-1}) - \Delta R(w_r)$  so inequality (14) holds. We have,  $\Delta H(v, w_r) \leq \Delta L(w_r)$  for all  $v \in V$  holds for all  $v \in V$  since  $H(x, w_r)$  is the only hitting time to  $w_r$  that increases; all others decrease or are constant. By Corollary 3.6, we have  $\Delta T_{\text{bestmix}} \geq 0$ .  $\blacksquare$

Surgery  $\mathcal{S}_4$  is the right-hand version of  $\mathcal{S}_3$ . However, if  $L(w_r) = R(w_r) + 1$  then applying  $\mathcal{S}_4$  would lead to  $L_{\tilde{G}}(w_r) = R_{\tilde{G}}(w_r)$  which indicates that  $\tilde{G}$  has one focus by Lemma 3.2(a). We choose to avoid this situation, So we require that  $L(w_r) > R(w_r) + 1$  and handle the case  $L(w_r) = R(w_r) + 1$  with  $\mathcal{S}_5$  below.

**Lemma 4.5** *Let  $G$  be a bifocal caterpillar with  $L(w_r) > R(w_r) + 1$  that contains distinct vertices  $y \in U_{t-1}$  and  $x \in U_i$  where  $r < i < t$ . If  $\mathcal{S}_4 = \tau(t-1, t-2) \wedge \tau(i, i-1)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** This surgery extends the spine on the right-hand side. Analogous to the previous proof, this time the left-pessimal hitting times have  $\Delta L(w_{r-1}) = 0 = \Delta L(w_r)$  and the right-pessimal hitting times have  $\Delta R(w_{r-1}) = 1 = \Delta R(w_r)$ . Inequality (13) holds because  $L(w_r) - R(w_r) > 1 = \Delta R(w_r) - \Delta L(w_r)$ , and inequality (14) holds because  $\Delta L(w_{r-1}) - \Delta R(w_r) < 0$ . By Lemma 3.3,  $\tilde{G}$  has focus  $w_{r-1}$  and best mix focus  $w_r$ .

Next, we show that  $\Delta T_{\text{bestmix}} > 0$  using Lemma 3.5 with  $A = \{y\}$ . We have  $\Delta L(w_r) < \Delta H(y, w_r) = \Delta R(w_r)$ , while  $\Delta H(v, w_r) \leq 0 = \Delta L(w_r)$  for  $v \neq y$ . The left hand side of inequality (15) equals  $\Delta H(y, w_r) = 1$ . We take  $B = \{w_{i-1}, w_i, w_{t-1}, w_t\}$  because the contribution from each vertex in  $\overline{A} \setminus B$  is positive. Using equation (10), this sum is

$$\begin{aligned} & -H_G(w_{i-1}, w_r) + (2 \deg_{\tilde{G}}(w_i) + H_G(w_i, w_r)) + (2 \deg_{\tilde{G}}(w_{t-1}) + H_G(w_{t-1}, w_r)) \\ & \quad + (-2 \deg_{\tilde{G}}(w_t) - H_G(w_t, w_r)) \\ & = H_G(w_i, w_{i-1}) - H_G(w_t, w_{t-1}) + 2 (\deg_{\tilde{G}}(w_i) + \deg_{\tilde{G}}(w_{t-1}) - \deg_{\tilde{G}}(w_t)) \\ & > H_G(w_i, w_{i-1}) > 1. \end{aligned}$$

By Lemma 3.5,  $\Delta T_{\text{bestmix}} > 0$ . ■

We now consider surgery  $\mathcal{S}_5$ , which is only used when  $L(w_r) = R(w_r) + 1$ . This is one of the crucial moments in our algorithm: a minor change threatens the balance described in Lemma 3.2. In fact, surgery  $\mathcal{S}_5 = \tau(t-1, t)$  is the first surgery that shifts the location of the foci. The resulting graph is bifocal with best mix focus  $w_r$ , but the second focus moves from  $w_{r-1}$  to  $w_{r+1}$ .

**Lemma 4.6** *Let  $G$  be a bifocal caterpillar with  $y \in U_{t-1}$  such that  $L(w_r) = R(w_r) + 1$ . If  $\mathcal{S}_5 = \tau(t-1, t)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

We note that during Phase One, we will also have a second right leaf (otherwise we are done with the right spine). However, our proof does not require the existence of such a leaf.

**Proof:** Let  $\tilde{G} = \mathcal{S}_5(G)$  and let  $y \in V_{t-1}$  be the transplanted vertex. The changes in the left and right hitting times to the foci are  $\Delta L(w_{r-1}) = 0 = \Delta L(w_r)$  and  $\Delta R(w_{r-1}) = 3 = \Delta R(w_r)$ . Crucially, inequality (13) is not satisfied. We now verify that  $w_r$  is the best mix focus of  $\tilde{G}$  and that  $w_{r+1}$  is the other focus. We use Lemma 3.2 (d), replacing  $r$  with  $r+1$  and swapping left for right. In other words we must show that  $R(w_r) > L(w_r)$  and  $R(w_{r+1}) \leq L(w_r)$ . Observe that  $R_{\tilde{G}}(w_r) = R_G(w_r) + 3 > L_G(w_r) = L_{\tilde{G}}(w_r)$ , and

$$R_{\tilde{G}}(w_{r+1}) = R_G(w_{r+1}) + 3 < R_G(w_{r+1}) + H_G(w_{r+1}, w_r) = R_G(w_r) < L_G(w_r) = L_{\tilde{G}}(w_r).$$

Next, we show that  $T_{\text{bestmix}}(\tilde{G}) > T_{\text{bestmix}}(G)$  using Lemma 3.5. Note that the surgery shifts the foci, so  $w_1$  is the  $G$ -pessimal vertex for  $w_r$ , while  $y$  is the  $G'$ -pessimal vertex for  $w_r$ . We find that  $\Delta\text{Pess}(w_r) = 2$  because

$$\text{Pess}_{\tilde{G}}(w_r) = H_{\tilde{G}}(y, w_r) = 3 + H_G(w_t, w_r) = 2 + H_G(w_1, w_r) = 2 + \text{Pess}_G(w_r).$$

We will use Lemma 3.5 with  $A = \{y\}$ . The left side of equation (15) is  $\deg_{\tilde{G}}(y)(\Delta H(y, w_r) - 2) = 3 - 2 = 1$ . We take  $B = \{y, w_{t-1}, w_t\}$  since  $v \notin B$  means that  $\Delta \deg(v) = 0$  and  $\Delta H(v, w_r) \leq 2 = \Delta\text{Pess}(w_r)$ . The right hand side of equation (15) is

$$(2 \deg_{\tilde{G}}(w_{t-1}) + H_G(w_{t-1}, w_r)) - H_G(w_t, w_r) = 2 \deg_{\tilde{G}}(w_{t-1}) - H_G(w_t, w_{t-1}) > 1,$$

so by Lemma 3.5,  $\Delta T_{\text{bestmix}} > 0$ . ■

## 4.2 Phase Two: Seesaw to Twin Broom

In this section, we discuss Phase Two, which inputs a seesaw and outputs a twin broom.

Surgery	Initial Conditions	Illustration
$\mathcal{S}_6$	$G$ has exactly one left leaf and $2r - 1 \leq t$	
$\mathcal{S}_7$	$2r - 2 \geq t$ ; $G$ has exactly one left leaf and one right leaf	
$\mathcal{S}_8$	$G$ has exactly one right leaf and no left leaves	

Table 4.2: The Phase Two tree surgeries. The tree  $G$  is a bifocal seesaw.

**Lemma 4.7** *Let  $G$  be a seesaw on  $n$  vertices. Phase Two creates a twin broom  $\tilde{G}$  such that  $T_{\text{bestmix}}(\tilde{G}) \geq T_{\text{bestmix}}(G)$ .*

**Proof:** If  $G$  is already a twin broom then  $\tilde{G} = G$ . The three surgery types  $\mathcal{S}_6, \mathcal{S}_7, \mathcal{S}_8$  employed in Phase Two are shown in Table 4.2 and the workflow is shown in Figure 4.3. The lemmas that follow show that  $\Delta T_{\text{bestmix}} \geq 0$  for all three surgeries.

Suppose that  $G$  has a left leaf. If  $2r - 1 \leq t$  then we use  $\mathcal{S}_6$  to transplant the left leaf onto the end of the left spine. If  $2r - 2 \geq t$ , or equivalently  $r - 2 \geq t - r$ , then there must also be a right leaf because  $w_r$  is the best mix focus (see Lemma 3.2(c)). In this case, we use  $\mathcal{S}_7$  to simultaneously transplant both of these leaves to the spine. Next, if  $G$  still has a right leaf (perhaps we just applied  $\mathcal{S}_6$ ), then we apply  $\mathcal{S}_8$  to move this leaf to the end of the right spine. The resulting graph is a twin broom with best mix focus  $w_r$  and second focus  $w_{r-1}, w_r$ . ■

First, we prove that  $\mathcal{S}_6$  moves the final left leaf to the spine while also increasing  $T_{\text{bestmix}}$ .

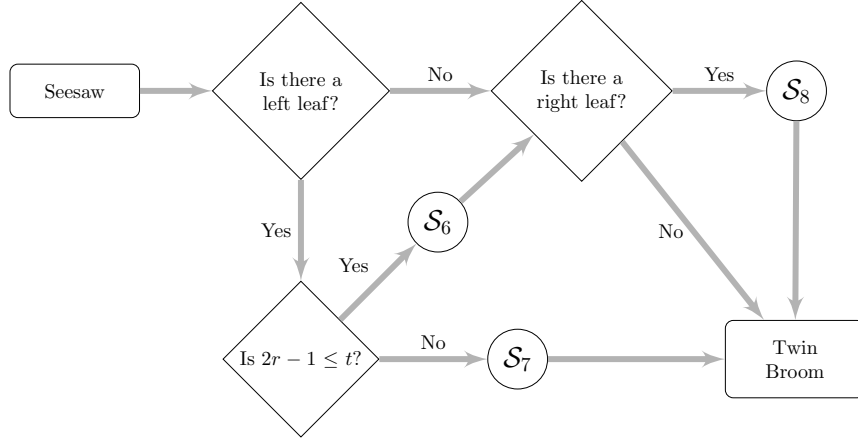


Figure 4.3: Phase Two of the algorithm, turning a seesaw into a twin broom.

**Lemma 4.8** *Let  $G$  be a seesaw with  $2r - 1 \leq t$  and exactly one left leaf  $x \in U_i$ , where  $2 \leq i \leq r - 2$ . If  $\mathcal{S}_6 = \tau(i, 1)$  then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** The surgery extends the left-hand spine, so  $L(w_{r-1}) = H_{\tilde{G}}(x, w_{r-1}) = H_{\tilde{G}}(x, w_1) + H_{\tilde{G}}(w_1, w_{r-1}) = 1 + H_G(w_1 + w_{r-1}) + 2(i - 1)$  by equation (7). Therefore  $\Delta L(w_{r-1}) = 2i - 1$ , and likewise  $\Delta L(w_r) = 2i - 1$ , while  $\Delta R(w_{r-1}) = 0 = \Delta R(w_r)$ .

We show that the foci have not changed using Lemma 3.2 (d). We have  $L_{\tilde{G}}(w_r) > L_G(w_r) \geq R_G(w_r) = R_{\tilde{G}}(w_r)$ . Equation (4) and the assumption  $2r - 1 \leq t$  yield  $L_{\tilde{G}}(w_{r-1}) = (r - 1)^2 \leq (t - r)^2 \leq R_{\tilde{G}}(w_r)$ . Therefore,  $\tilde{G}$  has focus  $w_{r-1}$  and best mix focus  $w_r$ .

Next, we use Lemma 3.5 with  $A = \{x\}$ . The left-hand side of inequality (15) is

$$\Delta H(x, w_r) - \Delta L(w_r) = (i^2 - 1) - (2i - 1) = i^2 - 2i.$$

Let  $B = \{w_{r-1}, w_r, \dots, w_{t-1}\}$ . For all  $v \in B$ , we have  $\deg_{\tilde{G}}(v) \geq 2$  and  $\Delta H(v, w_r) = 0 = \Delta \deg(v)$ . The right-hand side of inequality (15) is

$$\sum_{v \in B} 2(2i - 1) = 2(2i - 1)(t - r) \geq 2(2i - 1)(r - 2) > i(i - 2) = i^2 - 2i.$$

Lemma 3.5 gives  $\Delta T_{\text{bestmix}} > 0$ . ■

We employ surgery  $\mathcal{S}_7$  is the particular case where  $2r - 2 \geq t$  and  $G$  has both a left leaf and a right leaf. This surgery removes both leaves simultaneously.

**Lemma 4.9** *Let  $G$  be a seesaw with exactly one left leaf  $x \in U_i$  and exactly one right leaf  $y \in U_j$ , such that  $2r - 2 \geq t$ . If  $\mathcal{S}_7 = \tau(i, 1) \wedge \tau(j, t)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** First, we show that in fact,  $2r - 2 = t$ , meaning that the left and right spines are the same length. If  $r - 2 > t - r$  then  $L(w_{r-1}) = H(w_1, w_{r-1}) > (r - 2)^2 \geq (t - r + 1)^2 >$

$H(w_t, w_r) = R(w_r)$ , which contradicts Lemma 3.2(c) because  $w_r$  is a best mix focus. Next, we observe that  $i - 1 \geq t - j$ , meaning that  $i$  is further from the left endpoint than  $j$  is from the right endpoint. Indeed, when the left and right spines are the same length, this is necessary for  $L(w_{r-1}) \leq R(w_r)$ .

By equations (7) and (10), we have  $\Delta L(w_{r-1}) = 2i - 1 = \Delta L(w_r)$  and  $\Delta R(w_{r-1}) = 2(t - j) + 1 = \Delta R(w_r)$ . After the surgery, the left and right spines are equal length and leaf-free, so the conditions of Lemma 3.2(d) hold for  $\tilde{G}$ . In fact, both  $w_{r-1}, w_r$  are now best mix foci because  $L_{\tilde{G}}(w_{r-1}) = R_{\tilde{G}}(w_r)$ . To prove that  $T_{\text{bestmix}}$  increases, we use Lemma 3.5. Taking  $A = \{x, y\}$ , the left hand side of inequality (15) is

$$(i^2 - 1 - (2i - 1)) + ((t + 1 - j)^2 - 1 - (2i - 1)) \leq 2(i^2 - 1 - (2i - 1)) = 2i^2 - 4i.$$

For the right hand side of inequality (15), we take  $B = W = \{w_1, w_2, \dots, w_t\}$ , in other words, we ignore any central leaves. This right hand side is at least

$$\begin{aligned} & 2t \Delta \text{Pess}(w_r) - 2 \sum_{k=1}^t \Delta H(w_i, w_r) \\ & \quad - H_G(w_1, w_r) + H_G(w_i, w_r) + H_G(w_j, w_r) - H_G(w_t, w_r) \\ & = 2t(2i - 1) - 4 \sum_{k=1}^{i-1} (i - k) - 4 \sum_{k=j+1}^t (k - j) - (i - 1)^2 - (t - j)^2 \\ & \geq 2t(2i - 1) - 4i(i - 1) - 2(i - 1)^2 = 4(r - 1)(2i - 1) - 6i^2 + 8i - 2 \\ & > 4(i - 1)(2i - 1) - 6i^2 + 8i - 2 = 2i^2 - 4i + 2. \end{aligned}$$

We have satisfied the conditions of Lemma 3.5, so  $\Delta T_{\text{bestmix}} > 0$ . ■

Finally, we consider surgery  $\mathcal{S}_8$ , which moves the final right leaf to the spine.

**Lemma 4.10** *Let  $G$  be a seesaw with exactly one right leaf  $x \in U_i$ , where  $r + 1 \leq i \leq t - 1$  and no left leaves. If  $\mathcal{S}_8 = \tau(i, r)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** First, we use Lemma 3.3, to show that  $\tilde{G}$  has focus  $w_{r-1}$  and best mix focus  $w_r$ . By equation (9),  $\Delta L(w_{r-1}) = 0 = \Delta L(w_r) = \Delta H(w_1, w_r)$  and  $\Delta R(w_{r-1}) = -2(i - r) = \Delta R(w_r)$ . Inequality (13) is clearly satisfied. Next, we verify inequality (14). By equation (4), we have  $R_G(w_r) - L_G(w_{r-1}) = (t - r)^2 + 2(i - r) - (r - 2)^2$  and we claim that  $(t - r)^2 - (r - 2)^2 > 0$ . Indeed, since  $w_r$  is a best mix focus, Lemma 3.3 (c) ensures that  $L(w_{r-1}) \leq R(w_r)$ , or in other words  $(r - 2)^2 \leq (t - r)^2 + 2(r - i) < (t - r + 1)^2$ . Since  $r - 2$  and  $t - r$  are both integers, we must have  $r - 2 \leq t - r$ . This means that  $R_G(w_r) - L_G(w_{r-1}) \geq 2(i - r) = \Delta L(w_{r-1}) - \Delta R(w_r)$ . This confirms that the foci do not change. Finally, for all  $v \in V$ , we have  $\Delta H(v, w_r) \leq 0 = L(w_r)$ . By Corollary 3.6,  $\Delta T_{\text{bestmix}} > 0$ . ■

### 4.3 Phase Three: Twin Broom to Path or Wishbone

Phase Three converts a twin broom into either  $P_n$  or  $Y_n$ . The three surgeries are shown in Table 4.3 and the workflow is shown in Figure 4.4.

**Lemma 4.11** *Let  $G$  be a twin broom on  $n$  vertices. If  $n$  is even then Phase Three turns  $G$  into the path  $P_n$ . If  $n$  is odd, then Phase Three turns  $G$  into the wishbone  $Y_n$ . Moreover, if  $n$  is even then  $T_{\text{bestmix}}(G) \leq T_{\text{bestmix}}(P_n)$  and if  $n$  is odd then  $T_{\text{bestmix}}(G) \leq T_{\text{bestmix}}(Y_n)$ . Furthermore, equality holds if and only if  $G = P_n$  for  $n$  even, and  $G = Y_n$  for  $n$  odd.*

**Proof:** The Phase Three surgeries  $\mathcal{S}_9, \mathcal{S}_{10}, \mathcal{S}_{11}$  are shown in Table 4.3 and the workflow is shown in Figure 4.4. The lemmas that follow show that  $\Delta T_{\text{bestmix}} > 0$  for all three surgeries.

Suppose that  $r - 2 < t - r$ , or equivalently  $t < 2r - 2$ . Since  $G$  is bifocal, there must be enough leaves in  $U_{r-1}$  so that  $(r - 1)^2 + 2|U_{r-1}| = L(w_r) > R(w_r) = (t - r)^2$ . We apply  $\mathcal{S}_9$  until  $r - 2 = t - r$ . At this point (and henceforth), both  $w_{r-1}$  and  $w_r$  are best mix foci by Lemma 3.2 (d). We may assume that  $|U_{r-1}| \leq |U_r|$ . If both  $U_{r-1}$  and  $U_r$  are nonempty, we use  $\mathcal{S}_{10}$  to simultaneously extend the left and right spine by taking one vertex from each of these sets. We repeat this until  $U_{r-1} = \emptyset$ . Next, if there are multiple leaves remaining in  $U_r$ , we use  $\mathcal{S}_{11}$  to extend both ends of the spine. We repeat  $\mathcal{S}_{11}$  until there are 0 or 1 leaves left in  $U_r$ . At this point, we either have a path  $P_n$  or a wishbone  $Y_n$ . We started this phase with a twin broom, which is bifocal by definition. Therefore, if  $n$  is even we have constructed a path and if  $n$  is odd we have constructed a wishbone. ■

Surgery	Initial Conditions	Illustration
$\mathcal{S}_9$	$r - 2 < t - r$ , which forces $U_{r-1} \neq \emptyset$	
$\mathcal{S}_{10}$	$r - 2 = t - r$ , and there are leaves $x \in U_{r-1}$ and $y \in U_r$	
$\mathcal{S}_{11}$	$r - 2 = t - r$ , $U_{r-1} = \emptyset$ and there are leaves $x, y \in U_r$	

Table 4.3: The Phase Three tree surgeries. The graph  $G$  is a twin broom.

We first consider  $\mathcal{S}_9$ , which we apply to equalize the lengths of the left and right spine.

**Lemma 4.12** *Let  $G$  be a twin broom with leaf  $x \in U_{r-1}$  such that  $2r - 1 \leq t$ . If  $\mathcal{S}_9 = \tau(r - 1, 1)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** First, it is routine to show that the simultaneous inequalities  $2r - 1 \leq t$  and  $L_G(w_r) > R_G(w_r)$  require that  $\deg(w_{r-1}) \geq 4$ . By equation (7), we have  $\Delta L(w_{r-1}) = 2r - 3 = \Delta L(w_r)$

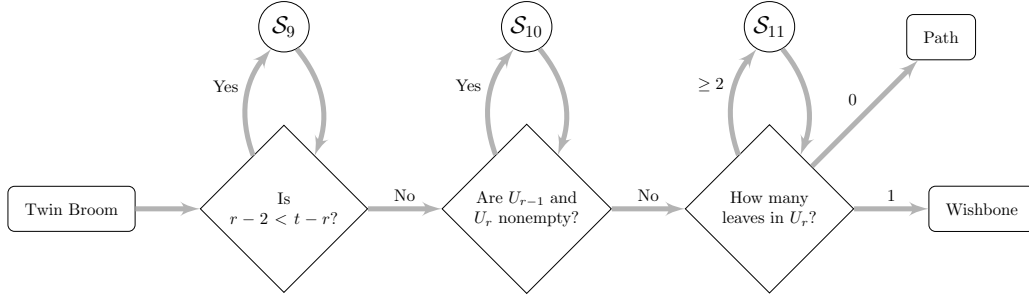


Figure 4.4: Phase three of the algorithm, turning a twin broom into a path or a wishbone.

and  $\Delta R(w_{r-1}) = 0 = \Delta R(w_r)$ . We use the two criteria of Lemma 3.2 (d) to show that the foci roles do not change. Clearly  $L_{\tilde{G}}(w_r) > L_G(w_r) \geq R_G(w_r) = R_{\tilde{G}}(w_r)$ . Since  $U_i = \emptyset$  for all  $i \notin \{r-1, r\}$ , we can use equation (4) to calculate  $L_{\tilde{G}}(w_{r-1}) = (r-1)^2 \leq (t-r)^2 = R_{\tilde{G}}(w_r)$ .

Next, we show that the best mixing time increases using Lemma 3.5 with  $A = \{x\}$ . The left-hand side of inequality (15) is

$$\Delta H(x, w_r) - \Delta \text{Pess}(w_r) = ((r-1)^2 - 1) - (2r-3) = r^2 - 4r + 3.$$

For the right hand side, we take  $B = W$ , ignoring any other central leaves. We obtain

$$\begin{aligned}
 & (2t-1) \Delta \text{Pess}(w_r) - 2 \sum_{i=1}^{r-1} \Delta H(w_i, w_r) - H_G(w_1, w_r) + H_G(w_{r-1}, w_r) \\
 &= (2t-1)(2r-3) - 4 \sum_{i=1}^{r-1} (r-1-k) - (r-2)^2 \\
 &\geq 2(2r-2)(2r-3) - 3r^2 + 10r - 8 = 5r^2 - 10r + 4,
 \end{aligned}$$

which is clearly larger than the left hand side, so  $\Delta T_{\text{bestmix}} > 0$ . ■

Next, we discuss transplanting one leaf from each of  $U_{r-1}$  and  $U_r$ .

**Lemma 4.13** *Let  $G$  be a twin broom where  $t = 2r - 2$  and with vertices  $x \in U_{r-1}$  and  $y \in U_r$ . If  $\mathcal{S}_{10} = \tau(r-1, 1) \wedge \tau(r, t)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** The right-hand and left-hand spines have equal lengths before and after this surgery. Corollary 3.4 holds because each of  $\Delta L(w_{r-1})$ ,  $\Delta L(w_r)$ ,  $\Delta R(w_{r-1})$ , and  $\Delta R(w_r)$  are equal to  $2(r-2) - 1 = 2r - 3$ . So  $\tilde{G}$  has focus  $w_{r-1}$  and best mix focus  $w_r$ . We now show that the best mixing time increases. Setting  $A = \{x, y\}$ , the left hand side of inequality (15) is

$$\Delta H(x, w_r) + \Delta H(y, w_r) - 2(2r-3) = 2((r-1)^2 - 1) - 2(2r-3) = 2r^2 - 8r + 6.$$

where we use equations (7) and (10). As for the right-hand side, we take  $B = W$ , disregarding the remaining central leaves. This right hand side is at least

$$\begin{aligned} 2t \Delta \text{Pess}(w_r) - 2 \sum_{k=1}^t \Delta H(w_k, w_r) - H_G(w_1, w_r) + H(w_{r-1}, w_r) - H_G(w_t, w_r) \\ = 2(2r-2)(2r-3) - 2 \cdot 2 \sum_{k=1}^{r-1} 2(r-1-k) - 2(r-2)^2 = 2r^2 - 4. \end{aligned}$$

By Lemma 3.5,  $\Delta T_{\text{bestmix}} > 0$ . ■

Finally, we discuss transplanting pairs of leaves from  $U_r$  to the ends of the spine.

**Lemma 4.14** *Let  $G$  be a twin broom with  $t = 2r-2$  where  $U_{r-1} = \emptyset$  and with distinct vertices  $x, y \in U_r$ . If  $\mathcal{S}_{11} = \tau(r, 1) \wedge \tau(r, t)$ , then  $\Delta T_{\text{bestmix}} > 0$ .*

**Proof:** This proof is similar to the previous one: the left and right spine lengths are equal before and after the surgery. We have  $\Delta L(w_r) = 2r-1$  while  $\Delta L(w_{r-1}) = \Delta R(w_{r-1}) = \Delta R(w_r) = 2r-3$ , where the value for  $\Delta R(w_{r-1})$  takes into account the removal of two leaves from  $U_r$ . We have  $\Delta L(w_{r-1}) - \Delta R(w_r) = 0$  and  $\Delta R(w_r) - \Delta L(w_r) = -2$ , so Corollary 3.4 ensures that  $\tilde{G}$  has focus  $w_{r-1}$  and best mix focus  $w_r$ .

Setting  $A = \{x, y\}$ , the left hand side of inequality (15) is

$$(r^2 - 1) + ((r-1)^2 - 1) - 2(2r-1) = 2r^2 - 2r + 1.$$

As for the right-hand side of inequality (15), we take  $B = W$  and obtain

$$\begin{aligned} 2t \Delta \text{Pess}(w_r) - 2 \sum_{k=1}^t \Delta H(w_k, w_r) - H_G(w_1, w_r) - H_G(w_t, w_r) \\ = 2(2r-2)(2r-1) - 2 \sum_{k=1}^{r-1} 2(r-k) - 2 \sum_{k=r+1}^{2r-2} 2(k-r) - (r-1)^2 - (r-2)^2 \\ = 2r^2 + 2r - 5. \end{aligned}$$

Thus Lemma 3.5 ensures that  $\Delta T_{\text{bestmix}} > 0$ . ■

## 5 Conclusion

We have characterized the tree structures on  $n$  vertices that minimize and maximize  $T_{\text{bestmix}} = \min_{v \in V} H(v, \pi)$ . The star  $S_n$  is the unique minimizing structure, but the maximization problem depends on the parity of  $n$ . For even  $n$ , the maximizing structure is the path  $P_n$ , and for odd  $n$ , it is the wishbone  $Y_n$ . It is a bit strange that the odd path is not the maximizing structure for  $T_{\text{bestmix}}$ . But all is not lost: we believe that  $P_n$  is the maximizing structure for a slightly different quantity. For any graph  $G$ , the *forget distribution*  $\mu$  is the distribution achieving



$\max_{v \in V} H(v, \mu) = \min_{\tau} \max_{v \in V} H(v, \tau)$ . Lovász and Winkler [10] shows that  $\mu$  is unique, and they give a general formula. For a tree  $G$ , the forget distribution is concentrated on its foci [3]. When  $G$  is focal,  $\mu$  is a singleton distribution on the unique focus. When  $G$  is bifocal,  $\mu$  is given by

$$\mu_u = \frac{H(v', v) - H(u', v)}{2|E|} \quad \text{and} \quad \mu_v = \frac{H(u', u) - H(v', u)}{2|E|}$$

where  $u, v$  are the foci of the tree. Instead of  $T_{\text{bestmix}} = \min_w H(w, \pi)$ , we could instead consider the similar quantity  $H(\mu, \pi)$ . It is easy to see that  $S_n$  minimizes  $H(\mu, \pi)$  among all trees on  $n$  vertices. We conjecture that  $P_n$  maximizes this quantity for both even and odd  $n$ . Indeed, letting  $G = P_n$  and  $\tilde{G} = Y_n$ , calculations show that for odd  $n$ , we have  $H_{P_n}(\mu, \pi) = H_{Y_n}(\tilde{\mu}, \tilde{\pi}) + (2n - 3)/(2n - 2)$ , and for even  $n$ , we have  $H_{P_n}(\mu, \pi) = H_{Y_n}(\tilde{\mu}, \tilde{\pi}) + 1/(2n - 2)$ . Our tree surgery methods should be a fruitful line of attack, though tracking the changes in  $H(\mu, \pi)$  will require a new set of lemmas. We leave this problem for future work.

## References

- [1] D. Aldous and J. Fill. Reversible Markov Chains and Random Walks on Graphs. <http://www.stat.berkeley.edu/~aldous/RWG/book.html>, accessed January 2014.
- [2] D. Aldous, L. Lovász, and P. Winkler. Mixing times for uniformly ergodic Markov chains. *Stoch. Process. Appl.*, 71:165–185, 1997.
- [3] A. Beveridge. Centers for random walks on trees. *SIAM J. Discrete Math.*, 23(1):300–318, 2009.
- [4] A Beveridge and M. Wang. Mixing times for random walks on trees. *Graph Combin.*, 29(4):757–772, 2013.
- [5] G. Brightwell and P. Winkler. Maximum hitting time for random walks on graphs. In *Proc. 22nd ACM Symp. on Theory of Comput.*, pages 369–378, 1990.
- [6] I. Dumitriu, P. Tetali, and P. Winkler. On playing golf with two balls. *SIAM J. Discrete Math.*, 16(4):604–615, 2003.
- [7] U. Feige. A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(4):433–438, 1995.
- [8] A. Georgakopoulos and S. Wagner. Hitting times, cover cost, and the Wiener index of a tree. Preprint.
- [9] L. Lovász and P. Winkler. Efficient stopping rules for Markov chains. In *Proc. 27th ACM Symp. Theory Comput.*, pages 76–82, 1995.
- [10] L. Lovász and P. Winkler. Reversal of Markov chains and the forget time. *Combin., Probab. Comput.*, 7:189–204, 1998.